PROBLEM STATEMENT

• A system can detect traffic sign with high accuracy and run on real-time (>30 fps).

• Ability to infer to its abstract type when seeing an unknown sign.

Approach

• Built a You Only Look Once version 2 (YOLOv2) for Traffic Sign Detection.

• Embedded prior knowledge of the dataset into the network through a hierarchical graph.

Methods

• Based on the standard YOLO-v2, we extended the network architecture by using light-weight CNN with higher input resolution [2] using Densely Connected [3] and Depth-wise Separable Conv. Layers [4].

• Objective Function:

\[ w_1 \mathcal{L}_{\text{localization}} + w_2 \mathcal{L}_{\text{obj.confidence}} + w_3 \mathcal{L}_{\text{classification}} \]

EXPERIMENTS

• Trained on both LISA + LISA Extension, open-sourced US Traffic Sign Detection Dataset.

• Designed a Soft-Max Tree for joining multiple Traffic Sign dataset

RESULTS

<table>
<thead>
<tr>
<th>Network</th>
<th>Resolution</th>
<th>Parameters</th>
<th>mAP (%)</th>
<th>Forward Pass (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original YOLOv2</td>
<td>608x608</td>
<td>50.8 M</td>
<td>76.8</td>
<td>36ms</td>
</tr>
<tr>
<td>Mobile-YOLOv2</td>
<td>960x960</td>
<td>6.78 M</td>
<td>82.8</td>
<td>20ms</td>
</tr>
<tr>
<td>Densely-YOLOv2</td>
<td>960x960</td>
<td>10.6 M</td>
<td>90.8</td>
<td>55ms</td>
</tr>
</tbody>
</table>

(*) mAP is calculated on actual labels, without abstract classes. This result is preliminary. Additional training might increase the result.

(**) Validated on NVIDIA GTX 1080.

FUTURE WORK

• Use Recurrent Neural Network to take advantage of temporal information.

• Train on larger dataset.

• Occlusion Detection.

ACKNOWLEDGEMENT

Thank you Dr. Christoph Mertz, Research Associate Jina Wang, and all RISS staff for providing extraordinary support, resources and mentorship to help me complete this project.